12

NOSC

NOSC / TR 108

AD A041601

NOSC / TR 108

Technical Report 108

TEMPORARY AUDITORY THRESHOLD SHIFTS AND PERFORMANCE DECREMENTS ASSOCIATED WITH 12 DAYS OF NOISE EXPOSURE

A field study in fire room spaces aboard USS ORISKANY

JC Webster and FG Henry

15 April 1977

Final report
August 1975 to December 1976

Approved for public release; distribution is unlimited

NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CALIFORNIA 92152

NAVAL OCEAN SYSTEMS CENTER, SAN DIEGO, CA 92152

AN ACTIVITY OF THE NAVAL MATERIAL COMMAND

RR GAVAZZI, CAPT, USN
Commander

HOWARD L BLOOD, PhD
Technical Director

ADMINISTRATIVE INFORMATION

Work was done from August 1975 to December 1976 under NAVSEA 62543N SF 43452702 Task 1357, NOSC 740051, by the Human Factors Division. The report was approved for publication 15 April 1977.

Assistance in audiometry was provided by RG Klumpp and PO Thompson, of NOSC; and P Neff, of Naval Hospital, San Diego. Experimental design assistance was provided by D Seals and RT Wilkinson, of Naval Health Research Center.

Released by
John Silva, Head
Human Factors Division

Under authority of
RH DuBois, Head
Information Systems Department

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: NOSC Technical Report 108 (TR 108)

2. GOVT ACCESSION NO.:

3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): TEMPORARY AUDITORY THRESHOLD SHIFTS AND PERFORMANCE DECREMENTS ASSOCIATED WITH 12 DAYS OF NOISE EXPOSURE, (A field study in fire room spaces aboard USS ORISKANY).

5. TYPE OF REPORT & PERIOD COVERED: Final, August 1975 to December 1976

6. PERFORMING ORG. REPORT NUMBER:

7. AUTHOR(s): JC Webster and FG Henry

8. CONTRACT OR GRANT NUMBER(s):

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Naval Ocean Systems Center, San Diego, CA 92152

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62543N, SF 43452702 Task 1357

11. CONTROLLING OFFICE NAME AND ADDRESS: Naval Ocean Systems Center, San Diego, CA 92152

12. REPORT DATE: 15 April 1977

13. NUMBER OF PAGES: 36

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution is unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Noise (sound) - Dosimetry

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Effects of continual 1-in-3 watches were evaluated over a 10-day period. Eight (of 39) men were found with observable threshold shifts.

Dosimeter results showed that the typical noise exposure in terms of equivalent steady-state noise level in the space where most measurements were made averaged 89.9 dBA with a standard deviation of 4.5, whereas physical measures averaged 91.6 dBA with a standard deviation of 7.9.

On a psychomotor four-choice serial reaction time test all subjects (in noise and quiet) deteriorated on the average when performance at minutes 4 and 5 (of a 5-minute test) was compared to performance the first 2 minutes
(CONT)

20. (CONT)

The "quiet" group deteriorated less than the protected or unprotected noise group. Similarly, all three groups showed improvement on the reaction time test near the end of the watch. There apparently was a learning factor.

Perceptual interference was measured by a modification of the Stroop color word test. Again, all groups improved their overall performance on the retest toward the end of a watch.

Implementation of educational and compliance techniques is recommended to get more people to wear hearing protection to help save their hearing, improve their psychomotor performance, and reduce their perceptual interference in noise.

OBJECTIVE

Determine the effects of noise exposure on hearing performance of engineering personnel standing 1-in-3 watches in spaces having steady-state noise levels of 94 (±7) dBA.

RESULTS

1. Effects of continual 1-in-3 watches (4 hours in noise and 8 out) over a 10-day period were evaluated. (See CONCLUSIONS in report proper.)

RECOMMENDATIONS

1. Verify the results of this study by repeating the most significant parts with an expanded shipboard team.
2. Develop a wearable noise level accumulator to supplement existing noise dosimeters and to better correlate performance and hearing tests to actual noise exposure.
3. Implement educational and compliance techniques to encourage more people to wear hearing protection to help save their hearing, improve their psychomotor performance, and reduce their perceptual interference in noise.

TABLE OF CONTENTS

INTRODUCTION

This report, one of very few field studies of the effects of noise on hearing and performance, is as important for its eventual shortcomings as for its solid scientific (statistically significant) findings. First, it should be pointed out that for budgetary reasons the original research proposal was significantly altered. As initially planned, it was to be a comprehensive study of the effects on engineering and air department personnel of noise, heat/humidity, ship routine, and oversea deployment stresses on hearing, performance, sleep, type and extent of sick call and morale. As actually carried out, the study was limited to the effects of noise exposure on hearing and performance of engineering personnel standing 1-in-3 watches in spaces having steady-state noise levels of 94 (±7) dBA. The original intent was to find the effects on recovery from noise-induced temporary threshold shifts of between-watch noise levels in berthing/messing spaces and recreation and/or routine activities. Initial plans were for doing the study on an aircraft carrier (CV) going into an extensive overhaul program so that noise control measures could have been instituted in nonwatch areas. Instead, operational constraints resulted in performing the study on a ship, the USS ORISKANY (CVA 34), that was making its last deployment before being decommissioned.

A second area of compromise concerned the original choice between a general field (ship) study and controlled laboratory experiment. Laboratory experimentation is appropriate for defining precise noise exposure limits, but field studies are necessary to validate them. If laboratory results are to be valid, differences between field and laboratory behavior of the subjects for the total day, week, or month must be accounted for. This involves elaborate simulations in the laboratory, as, for example, in Project PING, in which 20 sailors were confined to an ensonified barracks for 60 days (see ref 1, Cantrell, 1974), or gross extrapolations of data for example, performance test results for 30-minute exposures (see ref 2, Hartley and Adams, 1974) to predict results for 240 minutes. Present-day limitations on the use of human subjects and in fact the applicability of research results to operational situations (Mansfield Amendment) make justifying, setting up, getting subjects for, and running basic laboratory studies even more costly and time-consuming. For these reasons it is easier to justify field (ship) studies even though it is known in advance that maintaining adequate controls is difficult if not impossible. In this case the steady-state noise exposures involved (94 (±7) dBA) for an average of 8 hours per day (over a 3-day cycle) exceed allowable limits as specified by Occupational Safety and Health Administration Standards (OSHA) (ref 3) or US Navy (Bumed INST 6260.6B, ref 4) instructions. Therefore, justifying this risky exposure for a laboratory experiment would be time-consuming, expensive, and perhaps even futile, whereas USN machinist mates stand these types of watches day in and day out (at least on certain sustained steaming periods the type of condition chosen for this study). Similarly, the applicability of the results is again built into the study; these are precisely the conditions that now exist, in which men work and live. If properly controlled, this is a relatively inexpensive and time-saving method of avoiding a laboratory simulation.

---

[1] Cantrell, RW (1974), Prolonged exposure to intermittent noise: Audiometric, Biochemical, Motor, Psychological and Sleep Effects, Laryngoscope, Supplement 1, 84 (10), Part 2, October

[2] Hartley, LR, and Adams, RG (1974), Effect of noise on the Stroop test, J Exper Psych, 102 (1), 62-66

[3] US Department of Labor, Occupational Safety and Health Standards, Federal Register, XXXVII, Part II, Washington, DC, 1972

[4] US Navy, Bureau of Medicine Instruction 6260.6B, 5 March 1970

The more elaborate coordinated multidiscipline study originally planned and necessary for the understanding of the effects of noise on all measurable aspects of human endeavor will not be discussed here. But parts of the compromise, preplanned, carefully counterbalanced experimental design and its intended implementation will be described in order to illustrate the many problems inherent in field experiments as well as the compromises necessary to salvage useful data from them. As stated in the opening sentence, perhaps as much is to be learned from the problems encountered as from the results.

METHOD

EXPERIMENTAL GROUPS

There is a choice of two general methods of trying to control the measurements of the effects of noise exposure on experimental subjects: either use the same subjects under both noisy and quiet conditions, or use different groups, one in noise, the other in quiet. Since heat/humidity stresses were also usually present in the noisy areas studied here (fire/machinery rooms), a modification of the second method using three groups was employed. One group of 24 men, eight on each of three watches (8X3), was selected that stood watches in the noisy areas (94(±7) dBA) and by their own choice wore no hearing protection (see fig 1). This first group was called the Unprotected Noise (UN) group. A second group of 12 men (4X3) was selected that stood watch in the same noisy areas but who either habitually or for the duration of these tests wore hearing protectors plugs or muffs that attenuated sounds in the speech range, 500 4000 Hz, by 18 (±7) dB (see fig 2). This second group was called the Protected Noise (PN) group. A third group of 12, the Quiet (Q) group, stood watches in levels of 69 (±5) dBA (see fig 3).

From the very first baseline tests (while ORISKANY was still at dock in Alameda, California), and throughout the whole at-sea period, the preselected groupings were unavoidably subject to change. New men had to be added due to job and/or watch schedule rearrangements, sickness, absences without leave, confinement in the brig, and the reluctance of the volunteer subjects to continue as the experiment progressed.

Even the assignment of men to experimental groups had to be changed when it was noted that many members of the UN group wore sound-powered telephones (SPP) at their duty stations (see fig 4 and 5). Worse yet, some wore SPP part of the time but not always. That is, they occasionally substituted for, or swapped jobs for, one particular watch or part of it. For example, on the day they were given the performance tests, their duty included acting as an SPP talker (see fig 6). These men were consequently reclassified from the UN to the PN group when the results of the performance tests were interpreted. However, when asked at the termination of the experiment whether or not they wore protection, four of the subjects who were observed wearing SPP ear cushions said they were just pinch hitting for a mate on the day of the performance tests and were unprotected most of the time. In actuality, therefore, the composition of the sample populations varied somewhat for various phases of the study. For example, there were always 12 subjects in the Q group even though all 12 did not participate in every phase. For the audiometric tests there were 11 Q subjects, 11 in the PN group, 9 in the UN group, and 8 in the Assumed Unprotected group (AUN). For the two performance tests the subjects available were 12 Q, 15 PN, and 19 UN.

Figure 1. Subject in number 1 fire room who wore no hearing protection and was therefore in the unprotected noise (UN) group

Figure 2. Subject wearing ear muff in After Steering, one of the protected noise (PN) group

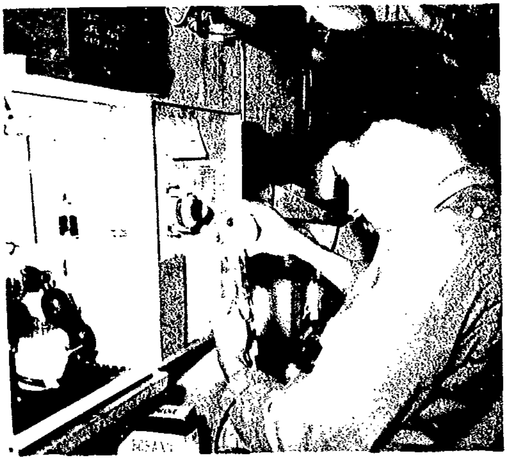

Figure 3. Subject in quiet (Q) group shown at his normal work station in the Oil Lab.

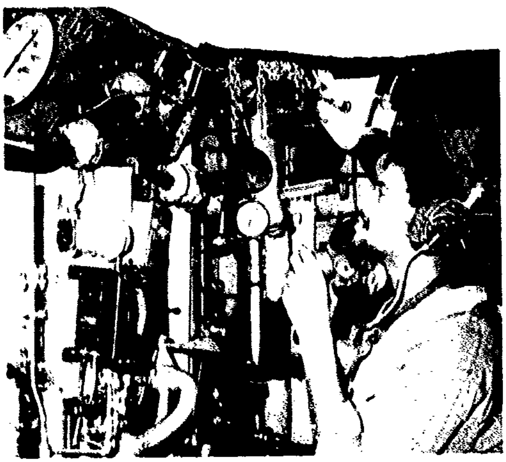

Figure 4. Subject originally assigned to UN group who at his duty station wore sound-powered telephone (SPP) muffs which partially protected his hearing. He was reclassified to the PN group in evaluating his performance tests.

Figure 5. Two subjects on the same watch in number 1 fire room. Subject at left was found to be wearing SPP so was reclassified to the PN group for the performance tests. Subject to right was in UN group.

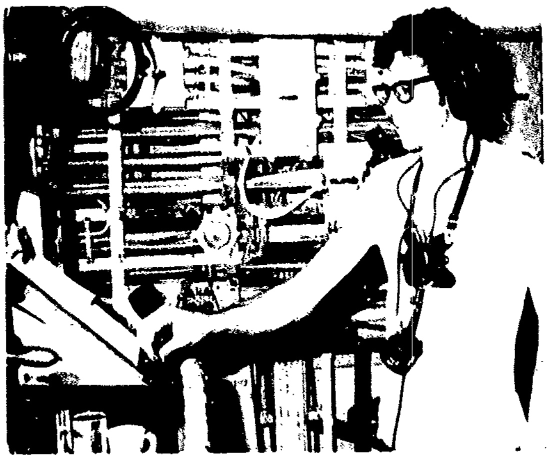

Figure 6. Subject in the UN group who, on occasion, performed a job which required him to wear a SPP headset, a form of hearing protection. This photograph, taken when the performance tests were given, places this man in the PN group for the performance tests but in the assumed unprotected noise (AUN) group for the audiometric tests.

NOISE EXPOSURE

Two factors determine noise exposure--noise level (at the inner ear) and duration of exposure. In practical terms this involves measuring the ambient noise level (at ear canal), the duration for each level, and the amount of sound attenuation (hearing protection) between the sound just outside the ear canal and the oval window (entrance to the inner ear). The original plans called for wearable noise dosimeters and sophisticated integrating and printout sound level equipments. The dosimeters were available, but the physical sound level equipments were tied up in another experiment, and the only available measuring equipment was a simple General Radio model 1565A hand-held survey sound level meter (SLM). A-weighted sound level measurements were taken from two to eight times at different hours on different days in each area in which subjects stood watches. These measurements were compared with those taken in the same spaces by US Navy Preventive Medicine Unit 6 a few months earlier under similar underway steaming conditions (ref 5).

The duration of noise exposure was ostensibly 4 hours (a normal watch period), followed by 8 hours in quieter places, repeated over a 2-week period. In order that the same men would not always stand the midnight to 0400 or 0400 to 0800 watch, a "dog watch" of 2 hours vice the usual 4 occurred from 1600 to 1800 and 1800 to 2000. A complete watch sequence repeated itself every 3 days; to wit, 0000-0400, 1200-1600, and 2000-2400; 0800-1200 and 1800-2000; and 0400-0800 and 1600-1800. Considering each day in turn, a man was noise-exposed 12 hours on day 1 and 6 hours on days 2 and 3, or an average of 8 hours per day. Of course, one-third of the subjects started their cycles on day 1, another third on day 2, and the final third on day 3. Some men who were not on watch per se, would nevertheless spend part of the usual working day (0800-1600) in the same noisy area in which they stood watch. On smaller ships this is often more the rule than the exception, but on ORISKANY, for the 2 weeks of the study, this was more the exception than the rule. Exact duration of exposure time when off watch was not, however, carefully accounted for. Attempts were made to account for off-watch activity by giving the subjects noise log cards to fill out, but they did not fill them out systematically or regularly and the attempt was given up.

To determine the exact noise exposures on watch, eight commercially available noise dosimeters were available to be worn by the subjects in this study. They were of three general types and two makes. Three made by Dupont were based on a 90-dBA threshold and the 5-dB doubling rule (5/90). In this study they were given the identity numbers 1, 2, and 3. Five were made by General Radio: number 4 was based on the 5/90 rule, numbers 5 and 6 were based on a 3/85 rule, and numbers 7 and 8 were based on a 3/60 rule. Figure 7 is a nomograph for converting the doses read out on these various rules to an equivalent steady-state noise level. For this study, this is a valid conversion, because the noise levels in the subject-occupied areas did not change appreciably in level either in time (over a 4-hour watch) or within the watch standing area. Some of the dosimeters had microphones on cords and clips and were worn at collar (ear) level, others had microphones integral with the unit and were worn at the chest pocket or held level. However, the noise levels were very constant both vertically and horizontally within these multireflective reverberant spaces. Microphone placement was not considered to be a problem of the same magnitude as other more identifiable problems and could certainly not be differentiated out with the meager data that were collected. Figure 8 shows the corpsman, RA Olson, HM1, calibrating and/or reading the dosimeters.

---

[5] US Navy, Environmental and Preventive Medicine Unit 6 encl (2) to ltr Ser EPMU6/6 vn. 6265 of 8 August 1975 to USS ORISKANY

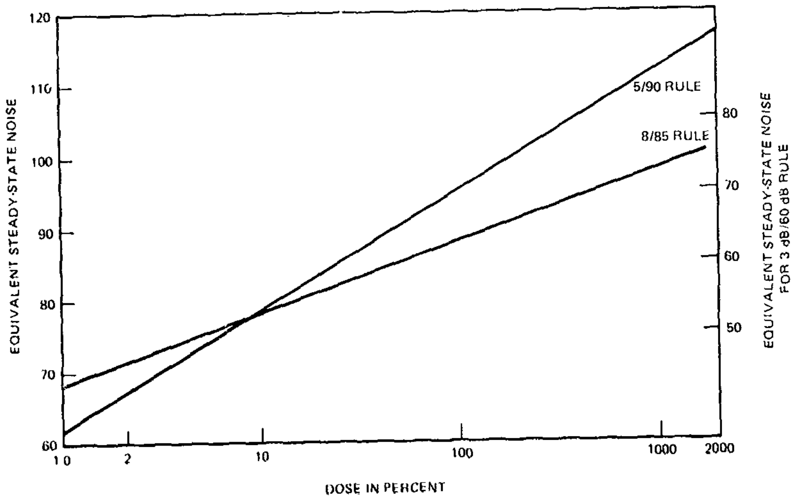

Figure 7 Chart for converting dose percentage to equivalent steady-state noise based on two rules: 90-dBA threshold and 5 dB per doubling (5/90) and 85-dBA threshold and 3 dB per doubling (3/85).

Figure 8 Corpsman calibrating dosimeters

PHYSIOLOGICAL AND PERFORMANCE TESTS

Three types of measures were made on noise-exposed and control subjects: audiometer tests, four-choice serial reaction times, and a modified version of the Stroop color word test. Each test is described in the following sections.

AUDIOMETER TESTS

The first concern of excessive noise exposure is damage to hearing, either a temporary threshold shift (TTS), which is readily measurable, or a permanent threshold shift (PTS), which is not so easily measurable. On the average, noise exposures of 94 dBA for 4 hours should produce some TTS if measured as prescribed; ie, within 2 minutes after cessation of noise exposure.

The audiometer used in these tests was the Grason-Stadler 1703 Recording Audiometer, a fixed-frequency Bekesy trace type audiometer that tests first at the frequencies 500, 1000, 2000, 3000, 4000, 6000, and 8000 Hz in one ear (the left), then at the same ordering of frequencies in the right ear, and finally at 1000 Hz in the right ear. Figure 9 shows two audiograms taken at 0850 and 1610 on a subject to show the expected consistency when tested in off-watch periods. Baseline audiograms were given to the 48 preselected men at dockside in Alameda, California, before ORISKANY left for Pearl Harbor en route to Subic Bay. These audiograms were given jointly by Dr RG Klumpp of NUC and Robert A Olson, HM1. Olson was being trained in the use of the 1703 by Dr Klumpp. Olson, or one of his colleagues, also trained by Dr Klumpp, gave all the audiometer tests once the ship was underway. More details on sweep speed, time on each frequency, etc, are available in the 1703 technical manual.

The background noise levels in the Industrial Acoustic Corporation (IAC) audiometric booth were checked out in detail by USN Preventive Medicine Unit 6 in August and spot checks were also made by Dr RG Klumpp before ORISKANY left Alameda. The booth met specifications. Figure 10 shows the audiometer, the booth, and Corpsman Olson. Figure 11 shows a man in the booth taking the audiometer tests.

Initial planning was centered on scheduling audiometer tests for four men at the beginning and the end of one of their five different 4-hour (but not their 2-hour) watches sometime during their 12-day steaming period. Figure 12 and table 1 show the initial plan. Note there are three watches covering a 12-day steaming period (from Pearl Harbor to Subic Bay in the Philippine Islands). The numerals above and below each 4-hour watch in figure 12 show the subsection scheduled for audiometer test on that particular watch. Note that in this manner each man (watch subsection) would be tested before and after each watch period once during the 12-day steaming period; to wit, and as shown in table 1, the first subsection of watch one (11) is tested before and after the 0000-0400 watch on day 1, the 0400-0800 watch on day 3, the 0800-1200 watch on day 5, the 2000-2400 watch on day 7 and the 1200-1600 watch on day 10. Table 1 details how all the other 15 subsections are covered in this same general manner.

A major objective of the experimental design was to get post-noise-exposure (end of watch) audiograms immediately. Consequently, within each group of subjects three-man teams were arranged, one from each watch, such that subject "a" (112 in table 2) could have a prewatch audiogram and replace "b" (111), who would then go immediately for his postwatch audiogram. Four hours later "c" (113) would have his pretest and replace

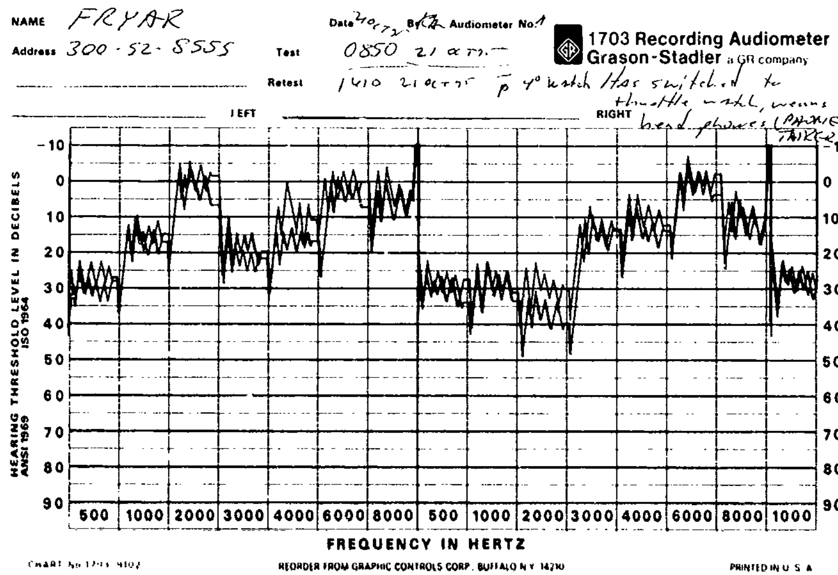

Figure 9. Sample audiogram.

"a" (1121), who would immediately have his post test. Four hours later "b" (1111) would have his pretest and replace "c" (1131), who would immediately have his posttest. This cycling would continue for the 12-day test regime. The rotation scheme was supplemented with a slight staggering of the watch periods so that the four 3-man control group teams (31 - -) would change watches 30 minutes before the scheduled watch change; the four 3-man PN teams (21 - -) would change 15 minutes before the watch; and the eight 3-man UN group teams (1x - -) would change watches on the hour (11 - -) and 15 minutes past the hour (12 - -).

At an initial meeting aboard ship before sailing, each man was given a 3 by 5 card listing the exact times he was to have prewatch and postwatch audiograms and to relieve and/or to be relieved by his other team members. The importance of having his postwatch audiogram immediately was emphasized repeatedly during this meeting.

In spite of initial briefings, the issuance of individual schedule cards, daily memos rotating the schedules to all involved engineering department divisions (A, B, E, M, and R) for posting, and the sending of messengers to round up delinquents, it was not possible to get the men to the audiometric test booth on schedule. It was a rare occasion indeed when even a single noise-exposed individual started his postwatch audiogram within 5 minutes of leaving the noise. Of the 10 scheduled audiometer tests for every subject (before and after each of the five different 4-hour periods scattered over 12 days), the average turnout was 50% and the average time after watch was 20 (±5) minutes. The delay was due to the

Figure 10. Corpsman administering audiometer test.

Figure 11. Subject inside audio booth taking audiometer test.

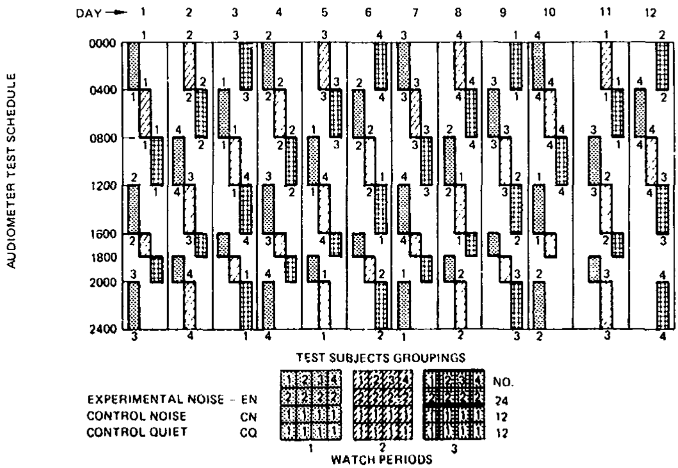

Figure 12. Twelve-day 1-in-3 watch schedule showing the scheduling of four subsections within a watch for audiometer tests (numerals above and below each 4-hour watch).

TABLE 1. AUDIOMETER TEST SCHEDULE FOR THE FIVE 4-HOUR WATCH PERIODS OVER THE 12-DAY CONTINUAL STEAMING PERIOD FOR THE FOUR SUBSECTIONS (SECOND NUMERAL) WITHIN EACH WATCH SECTION (FIRST NUMERAL).

		WATCH PERIOD				
		0/4	4/8	8/12	12/16	20/24
	1	1 1	2 1	3 1	1 2	1 3
	2	2 2	3 2	1 4	2 3	2 4
Day	3	3 3	1 1	2 1	3 4	3 1
	4	1 2	2 2	3 2	1 3	1 4
of	5	2 3	3 3	1 1	2 4	2 1
	6	3 4	1 2	2 2	3 1	3 2
Test	7	1 3	2 3	3 3	1 4	1 1
	8	2 4	3 4	1 2	2 1	2 2
	9	3 1	1 3	2 3	3 2	3 3
	10	1 4	2 4	3 4	1 1	1 2
	11	2 1	3 1	1 3	2 2	2 3
	12	3 2	1 4	2 4	3 3	3 4

TABLE 2. CODE NUMBERS FOR THE 48 EXPERIMENTAL SUBJECTS ARRANGED TO SHOW FROM TOP TO BOTTOM THE 16 THREE-MAN TEAMS WHO CONTINUALLY REPLACED EACH OTHER ON SUCCEEDING WATCHES. IN COLUMN 1 ARE LISTED THE THREE EXPERIMENTAL GROUPS, IN COLUMN 2 THE INDIVIDUAL SUBJECT IDENTIFIERS, IN COLUMN 3 THE WATCH SECTION NUMBER, AND IN COLUMN 4 THE SUBSECTION WATCH SECTION NUMBER.

Team		Individual Identifier 1	Individual Identifier 2	Individual Identifier 3	WATCH SUB-SECTION
1	UN	1 1 1 1	1 1 2 1	1 1 3 1	1
2	UN	1 2 1 1	1 2 2 1	1 2 3 1	1
3	PN	2 1 1 1	2 1 2 1	2 1 3 1	1
4	Q	3 1 1 1	3 1 2 1	3 1 3 1	1
5	UN	1 1 1 2	1 1 2 2	1 1 3 2	2
6	UN	1 2 1 2	1 2 2 2	1 2 3 2	2
7	PN	2 1 1 2	2 1 2 2	2 1 3 2	2
8	Q	3 1 1 2	3 1 2 2	3 1 3 2	2
9	UN	1 1 1 3	1 1 2 3	1 1 3 3	3
10	UN	1 2 1 3	1 2 2 3	1 2 3 3	3
11	PN	2 1 1 3	2 1 2 3	2 1 3 3	3
12	Q	3 1 1 3	3 1 2 3	3 1 3 3	3
13	UN	1 1 1 4	1 1 2 4	1 1 3 4	4
14	UN	1 2 1 4	1 2 2 4	1 2 3 4	4
15	PN	2 1 1 4	2 1 2 4	2 1 3 4	4
16	Q	3 1 1 4	3 1 2 4	3 1 3 4	4

queue at the testing booth, since everyone tended to change watches at the scheduled time instead of on the experimental staggered schedule. Detailed examinations of available prewatch and postwatch audiograms to identify TTS revealed none.

Consequently, a new diagnostic strategy was devised. Some time after 10 days at sea, two audiograms taken before and after a watch were compared to the subject's baseline audiogram taken after days and weeks in port (at Alameda prior to departure). Eight of 39 audiograms were judged by at least two people with audiological training to show some indication of TTS.

FOUR-CHOICE SERIAL REACTION TIME TESTS

Wilkinson (ref 6, 1969), in reviewing multiple-choice serial reaction time (RT) tests, found them to be sensitive indices of the effects of environmental stress. Wilkinson and Houghton (ref 7, 1975) designed a portable four-choice RT tester based on a battery-operated

[6]Wilkinson, RT (1969), Some factors influencing the effect of environmental stresses on performance, Psych Bull 72, 260-272

[7]Wilkinson, RT, and Houghton, D (1975), Portable Four-Choice Reaction Time Test with Magnetic Tape Memory, Behavior Research Methods and Instrumentation, 7 (5), 441 - 446, September

cassette tape recorder. The equipment is shown in figures 13 and 14. The subject's task is to push the switch that corresponds to the one lighted bulb selected at random by the device. If correct, a 2000-Hz tone is recorded on the tape. If incorrect, a 4000-Hz tone is recorded. The difference in time between the lighting of the red bulb and the switch press is the measure of RT. Whether correct or incorrect, the pressing of the switch turns off the bulb and triggers the next random choice of lighted bulb. Computer-aided readouts of cassettes containing data on tests given at the beginning and ending of the watch periods to a limited sample from all experimental groups listed the number of attempts, average RT, errors, and gaps (RTs > 1 second) at the end of each minute of the 5-minute test.

After other testing routines had been stabilized, subjects on the 0800-1200 and 1200-1600 watches were tested on the Wilkinson-Houghton device as soon as possible after commencing a watch, and the same subject was tested again toward the end of the watch. From one to four subjects were tested during a given watch so that the initial time in noise (or quiet) before the start-of-watch test varied from about 10 minutes to 40 minutes. The time on watch, between the first and second tests, averaged 3 hours ±15 minutes. No predeparture baseline tests—in fact, no practice tests of any type—were given. However, practice runs for as long as each subject desired were given prior to the actual timed 5-minute test.

MODIFIED STROOP COLOR WORD TEST

During the time that one experimenter was administering the four-choice RT test, the second experimenter was administering a form of the Stroop test to another subject in the work environment. The Stroop test measures perceptual interference. In the original (its more usual) form, five colors are used in three ways: color names in black ink (WORDS), color patches (HUES), and color names in incompatible hues of ink (for example, the word RED spelled in green ink (MIXED)). The subject's usual task is to name as many WORDS, HUES, or MIXED hues (not spelled words) as possible in a fixed time, or to sort out sets of labeled cards while an experimenter times him. The measure of perceptual interference is the ratio of (or differences between) ink hue (MIXED) scores and either one of the other two scores.

The testing materials available aboard ship (left over from project PING) consisted of only three colors (RED, GREEN, and BLUE), and initial testing showed this to be too easy a task, so color pairs were used as follows: pairs of color words, color patches, and color names printed in incompatible hues were placed on 36 cards. For the color names printed in black ink (WORDS) and for the color patches (HUES), there were four replications of all possible combinations; for example, there were four RED REDS, four RED BLUES, four RED GREENS, four BLUE BLUES, four BLUE REDS, four BLUE GREENS, four GREEN GREENS, four GREEN REDS, and four GREEN BLUES. For the incompatibly colored ink names (MIXED), there were necessarily four permutations of each combination; for example, RED RED was printed in hues of BLUE BLUE, BLUE GREEN, GREEN GREEN, and GREEN BLUE, as were RED BLUE, RED GREEN, etc. A box was made up with nine bins identified with pairs of colored patches (no printed words) and subjects were asked to sort the shuffled deck of 36 printed WORD pairs into the nine bins (see fig 15a and 15b). They were timed with a stop watch. Second, they sorted the 36 color patch cards (HUES) into the nine bins and, finally, the 36 incompatible hue word card pairs (MIXED). This was done on 35 subjects at the beginning and toward the end of a 4-hour watch period.

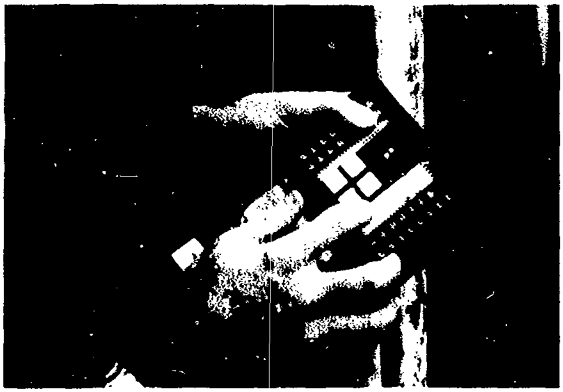

Figure 13. Picture of the Wilkinson-Houghton portable four-choice serial reaction time (RT) tester.

Figure 14. A subject protected by an SPP earphone cushion taking the RT test while on watch.

a.

b.

Figure 15. Pictures of modified Stroop color word test as used on ORISKANY.

always in the order WORDS, HUES, MIXED. It was done in the watchstanders' working environment. Since as many as four subjects were sometimes tested by one experimenter, the elapsed time between tests was closer to 3 hours than 4.

The basic raw score consists of the time in seconds required to completely sort the decks of cards into the bins, one score for the printed WORD pair in black ink, one for the color-patch (HUE) pairs, and one for the incompatible ink-hue word-pair cards (MIXED). There were very few errors, but when an error occurred a 3-second penalty was added onto the time score.

RESULTS AND DISCUSSION

AUDIOMETRY

As stated in Method, despite every effort to get the men to their postwatch audiometric test as soon as possible after their exposure to noise, it did not work out. Preliminary investigations of prewatch and postwatch audiograms showed a slight, nonsignificant increase in hearing acuity after the watch. Although not large enough to belabor, the apparent reason was increased arousal level. On average, and especially for those going on watch at midnight and 0400, and to a lesser degree for those going on at 0800 and 2000, the subjects were often half-asleep during the administration of the prewatch audiogram. At the conclusion of their watches, the men may have been physically tired, but they were certainly wide awake and apparently attended somewhat better to their audiometric tests.

Since the temporary threshold shifts at the conclusion of the 4-hour watches were not immediate and therefore of no diagnostic value, an effort was made to find whether anyone sustained a measurable hearing loss over a 10-day continual steaming period. Two audiograms on each man, one prewatch and one postwatch, were taken after at least 10 days at sea. These audiograms were then compared to each man's baseline audiogram. The resulting number of complete sets of three audiograms per person actually totaled 40. However, the swings or deflections made by one subject between his heard and not-heard judgments were often so long (±10 dB) that threshold level could not be determined objectively, and his audiometric set was discarded. Thus, the final total was 39 sets of audiograms. Of these, 11 were from the quiet group (Q, fig 16), 11 from the protected group (PN, fig 17), and 17 from the unprotected group. Of the unprotected 17, nine were known to be unprotected all the time (UN, fig 18). Some doubt existed about eight as to whether they did or did not wear earphones (protection) most of the time. These eight were removed from the unprotected group to make up a fourth group called assumed unprotected (AUN, fig 19).

The hearing levels (HLs) below 2 kHz for all groups were high, as if low-frequency noise masking was a factor. However, noise measurements in the booth indicated the contrary, leaving two possibilities: unsatisfactory earmuff seals or an out-of-calibration audiometer. Since the audiometer was calibrated before the study began, the earmuff seals must have been faulty.

The averages shown in figures 16 through 19 indicate higher baseline thresholds in the left ear at the three lowest frequencies. This could very well be a practice effect, since the left ear baseline thresholds were the first that a subject worked on in this testing format. This would explain the fact that the unprotected and assumed unprotected left ears showed no threshold shifts in the lower frequencies in the prewatch and postwatch audiograms while the right ears did show some, particularly in the postwatch tests.

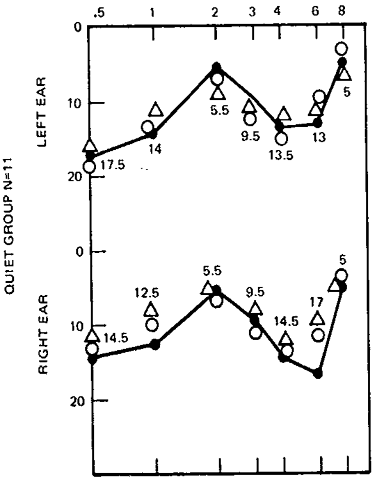

Figure 16. Baseline (●–●), prewatch (○), and post-watch (△) average hearing levels for 11 subjects in the Q group for the left ear (top) and the right ear (bottom).

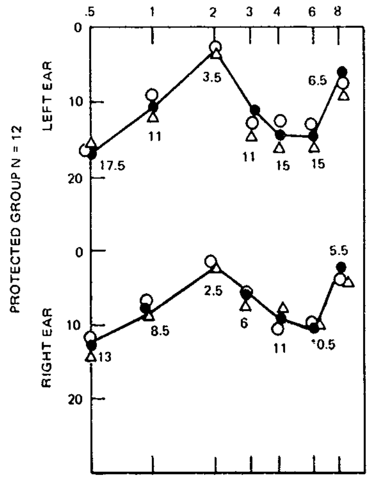

Figure 17. Baseline (●–●), prewatch (○), and post-watch (△) average hearing levels for 11 subjects in the PN group for the left ear (top) and right ear (bottom).

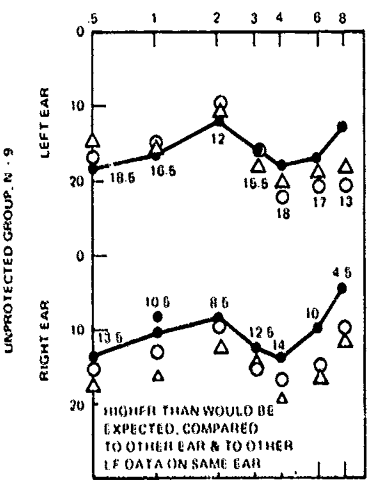

Figure 18. Baseline (●–●), prewatch (○), and post-watch (△) average hearing levels for 9 subjects in the UN group for the left ear (top) and right ear (bottom).

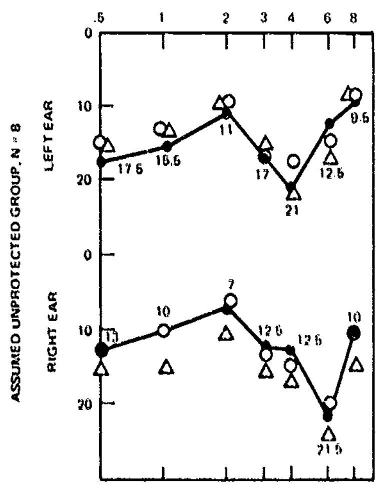

Figure 19. Baseline (●–●), prewatch (○), and post-watch (△) average hearing levels for 8 subjects in the AUN group for the left ear (top) and the right ear (bottom).

The high hearing levels (HLs) for all groups at 3, 4, and 6 kHz, shown in table 3, indicate that at least some members of each group have reduced acuity (probably PTS from long-term noise exposure). The quiet group shows no threshold shift in the postwatch audiograms, whether compared to prewatch test or baseline test. The protected group shows a mean 2.1-dB threshold shift in the postwatch test above 2 kHz in the left ears but none in the right ears. The assumed unprotected group shows a possible postwatch audiogram shift at 6 kHz in the left ears and at all frequencies in the right ears. The unprotected group shows shifts at all frequencies above 2 kHz in the left ears, as compared to baseline but not as compared to the prewatch audiogram; the right ears show the same at frequencies above 2 kHz, but there are also shifts at 1 and 2 kHz with respect to both baseline test and prewatch test. All postwatch 3–8-kHz threshold shifts from baseline audiograms were evaluated by the t-test of significance; the only one found to be statistically significant was the 5-dB mean shift at 4 kHz found in the right ears of the unprotected group. This shift was significant at better than 0.01 probability level. Significance of mean shifts of combined prewatch and postwatch thresholds from baseline was also checked (both by t-test, analysis of variance, and F-test), but these data did not reach the 0.01 level of significance.

As might be expected, the mean audiograms for all test sessions combined dipped lower in the high frequencies for the two unprotected groups than for the protected groups. The mean 3–8-kHz hearing levels were 10.4, 10.0, 15.3, and 15.5 dB for the quiet, protected, assumed unprotected, and unprotected groups, respectively. How much of this average 5-dB difference can be attributed to the particular noise environments described in this report is open to question, but the difference strengthens the possibility that there is a gradual occupational threshold shift accruing, though too gradual to measure over one steaming period with the degree of audiometric schedule deviations described.

Independent evaluations of the senior author and a research audiologist from the San Diego Naval Regional Health Center (Palmer Neff) found eight of the 39 men with observable changes in hearing level. Of these, four were in the unprotected noise group and three were in the assumed unprotected group, and one was from the quiet group who stood watches in the Engineering Log Room. Conversations with him (and some of his mates from the same berthing area) revealed that during any available leisure time he listened constantly to loud rock and roll music. His mates finally persuaded him to listen through earphones to keep the nuisance value down. He stated he always listened at the loudest level he could tolerate. Except for his left ear at 500 and 1000 Hz, he appears to have a rather consistent TTS of about 10 dB.

NOISE EXPOSURE

Although more detailed measurements are available, the general results can be summarized by noting that all spaces occupied by the noise-exposed subjects (1) were very stable in A-weighted noise level throughout any 4-hour watch at about 92 dBA (standard deviation = 7.9), (2) were consistent within any watchstanding area, and (3) were never outside the limits of 94 (±7) dB A-weighting, with C-weighting minus A-weighting levels of 5(±2) dB. The ambient level for the control subjects was 68(±5) dBA with C-A difference of 5(±2) dB. The more detailed measurements are discussed further in the section on dosimetry which follows.

TABLE 3. AVERAGE HEARING LEVELS (HL) FOR FOUR GROUPS AT 3, 4, 6, AND 8 kHz IN LEFT AND RIGHT EARS FOR BASELINE, AND PREWATCH AND POSTWATCH CONDITION.

Quiet group (N=11)

Left	Base	Pre-	Post-	Right	Base	Pre-	Post-
3 kHz	9.6	11.6	10.9	3 kHz	9.4	9.8	8.9
4	13.6	14.9	12.4	4	14.6	14.3	13.4
6	13.2	11.8	12.4	6	17 0	11.4	9.7
8	5.2	4.7	5.6	8	5.1	4.4	4.9
Mean	10.4	10.8	10.3	Mean	11.5	10.0	9.2

Mean HL for all conditions = 10.4

Protected

Left (N=11)				Right			
3 kHz	11.1	12.3	13.4	3 kHz	6.1	6.0	7.8
4	14.9	13.8	16.4	4	11.3	10.4	8.6
6	15.2	13.7	17.4	6	10.6	10.1	10.1
8	6.4	7.2	8.1	8	2.6	3.6	3.7
Mean	11.9	11.8	13.8	Mean	7.7	7.5	7.6

Mean HL for all conditions = 10.0

Assumed Unprotected

Left (N=8)				Right			
3 kHz	16.8	16.5	15.9	3 kHz	12.4	13.0	15.1
4	21.0	17.0	21.5	4	12.6	17.5	15.9
6	12.4	14.9	16.6	6	21.6	20.4	23.9
8	9.5	9.1	8.9	8	10.2	10.1	14.8
Mean	14.9	14.4	15.7	Mean	14.2	15.2	17.4

Mean HL for all conditions = 15.3

Unprotected

Left (N=9)				Right			
3 kHz	15.6	15.6	17.6	3 kHz	12.6	14.9	14.2
4	18.1	21.9	20.2	4	14.0	16.8	19.0
6	17.1	20.1	18.8	6	9.9	14.6	15.3
8	12.9	20 1	17.9	8	4.6	10.0	10.9
Mean	15.9	19.4	18.6	Mean	10.3	14.1	14.9

Mean HL for all conditions = 15.5

DOSIMETRY

Figure 20 shows the equivalent noise levels as measured by dosimeters which were worn on 77 separate occasions (4-hour watches) by 40 subjects in 11 different spaces. On 14 occasions the same man wore a 5/90 and a 3/85 rule dosimeter and these pairings are plotted according to the identity of the pairs as open symbols in figure 20. If (1) each member of the pair was accurately designed and calibrated, (2) the microphones were worn at the same general location, in breast pockets or on the belt, for the entire time, and (3) the assumption is correct that equivalent steady-state levels can be used as a datum, then these paired dosimeters should define a 45-degree line in figure 20. Of the 14 pairs, 11 are within 5 dB of defining the same noise level. The pairs at 70-78 (2 and 5), 83-68 (3 and 6), and 92-98 (4 and 7) do not meet these criteria. None of the out-of-line pairs have enough other points to determine whether they are in fact out of calibration or whether some other artifact entered into the mismatch. The pairs most often matched (1 and 5 six times and 2 and 6 three times) do fall into a linear configuration with a slope somewhat less than 45 degrees. The other points on figure 20 are measurements by a single 5/90 (x), 3/85 (+), or 3/60 (+) dosimeter and plotted on the 45-degree lines. They define a histogram of the noise exposures encountered in the study. There is some question of the validity of the four non-circled data points at 71, 72, and 73 dBA since they represent doses of 2, 2, 3, and 4 percent, which is also the case for the lower reading of the 70-78 and 83-68 pairs previously mentioned as being beyond the usable error range. Percentages of 4 or less are suspected to be due to circuitry noise (errors).

The modification of the 3/85 dosimeters to lower their thresholds to 60 dBA, called 3/60, was not satisfactory for usage in quiet areas aboard ships. The reason is that their dose reads only to 999, which corresponds to an equivalent steady-state level of 73 dBA. This obviously is not a sufficient range (note the truncation or bunching of datum points at 73 dB). Aboard aircraft carriers, an obviously better choice, or simple modification, would have been a threshold of 70 dB. In actuality, for research purposes, dosimeters that give the amount of accumulated time that the levels exceed each 3- or 5-dB step would be more appropriate.

The object of using the dosimeters was not to cross calibrate them, but to see whether they predicted in any way observed noise-induced hearing losses and/or to see whether their measures correlated with physical levels measured in the work areas. As stated previously, the prewatch versus postwatch audiometric data were not well controlled, but a measure of 10-day TTS was used to identify subjects. Tables 4 and 5 are laid out to answer both these questions: Are the noise exposures greater for the men showing some TTS? And how noisy are the work areas in terms of equivalent steady-state noise level (ESSNL)?

From table 4 it is apparent that there is no significant noise exposure level difference between the eight subjects who exhibited some 10-day TTS and those who did not. The only thing thoroughly convincing is that the subjects who normally wore hearing protection were <u>not</u> among the eight exhibiting TTS.

Concerning the comparison between ESSNL and physical noise levels, it can be observed from tables 4 and 5 that the agreement is within 2 dB and not statistically different. And both agree, and observation confirms, the levels in After Steering are generally higher and more variable.

Concerning doses within the same work area, there are variations (from 84 to 95 in Fire Room 1, 80 to 95 in Fire Room 2, 81 to 93 in Forward Engine Room, and 83 to 101 in After Steering). Physical measurements agree in general, but people do occasionally

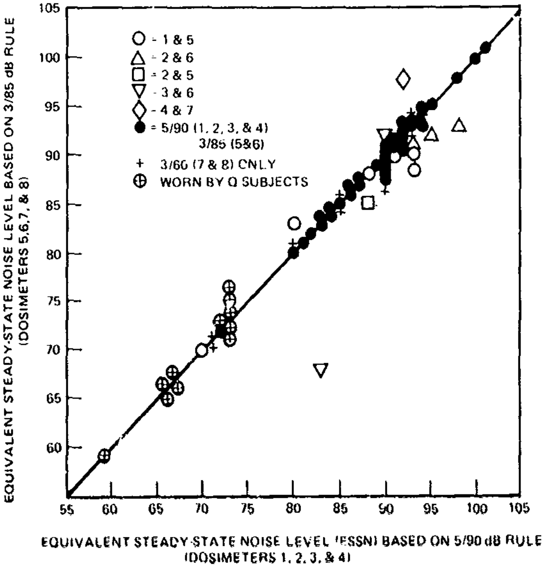

Figure 20. Equivalent steady-state noise levels for 53 dosimeter readings converted by using figure 7.

leave the noisy area and thereby reduce their exposure. If there were no differences, there would be no need for dosimeters in spaces such as these, which are in fact remarkably steady in noise level.

There are only two dosimeter readings in berthing areas (off watch), and both show levels well above those in the houses ashore in which it is assumed civilian industrial workers recover from their 8-hour noise exposures. In the original plans for this study the determination of leisure-hour space noise exposure levels was of major concern; but when the recording integrating sound level measurement equipment got tied up in a higher-priority and better-financed project, this phase had to be dropped. With the meager data available, the question of whether the eight subjects with 10-day TTS were more vulnerable to their noise environment during 4 hours on watch or to their lack of a quiet recovery area could not be answered.

FOUR-CHOICE SERIAL RT TEST

The results of the four-choice reaction test given at the beginning and end of a 4-hour watch are displayed in figures 21-23. Figure 21 shows the reaction time in seconds at the end of each one of the 5 minutes of the test. The figure is laid out in three groups, from top to bottom, eight of the noise-exposed unprotected subjects (UN), eight of the noise-exposed protected subjects (PN), and seven of the nonexposed control group (Q). The subjects are

TABLE 4. EQUIVALENT STEADY-STATE NOISE LEVELS (ESSNL) FROM DOSIMETER READINGS (FIG 20 FOR SELECTED WORK AND BERTHING AREAS. NUMBERS IN ITALICS ARE FROM DOSIMETERS WORN BY SUBJECTS WITH SOME 10-DAY TTS (SEE FIGS 16 - 19). NUMBERS IN BRACKETS ARE FROM PAIRS OF DOSIMETERS WORN BY THE SAME SUBJECT AT THE TIME OF MEASUREMENT. NUMBERS IN BOXES ARE FROM SUBJECTS WHO NORMALLY WORE HEARING PROTECTION.

SUBJECT WORK AREAS

	PROTECTED AND UNPROTECTED								QUIET				
	#1 Fire Room	#2 Fire Room	#3 Fire Room	#4 Fire Room	Forward Ending Room	Lower Level	After Steering	Average	Damage Control	Log Room	Oil Lab	Smoke Watch	Berth-ing
	93	95	87	80	91	98	101		67	≥73	≥73	≥73	≥73
	92	93			90	100	98		66	72	≥73	72	82
	92	89			91		93		66	≥73	66	59	
	92	90			88		95			70			
	92	93			85		94						
	91	91			93		93						
	90	92			93		92						
	90	91			91		87						
	85	90			90		84						
	84	90			89		83						
	95	90			88		83						
	92	90			88								
	93	83			86								
	93	80			86								
	90				83								
	89				81								
	85												
Mn	90.1	91.5	87.0		89.0	98.0		90.4		72.5			82.0
σ	3.1	2.8	0.0		2.5	0.0		3.3		0.7			0.0
N	10	4	1		5	1		21		2			1
Mn	91.0	89.0		80.0	88.0	100.0	91.1	89.7	66.7	71.5	70.7	68.0	73.0
σ	3.3	4.1		0.0	3.8	0.0	6.5	5.0	0.6	2.1	4.0	7.8	0.0
N	7	10		1	11	1	11	41	3	2	3	3	1
Mn	90.5	89.8	87.0	80.0	88.3	99.0	91.1	89.9	66.7	72.0	70.7	68.0	77.5
σ	3.1	3.9	0.0	0.0	3.4	1.4	6.5	4.5	0.6	1.4	4.0	7.8	6.4
N	17	14	1	1	16	2	11	62	3	4	3	3	2

TABLE 5. NOISE LEVELS IN MACHINERY SPACES.
EQUIVALENT STEADY-STATE NOISE LEVELS (ESSNL) BASED ON DOSIMETER READINGS AND PHYSICAL SOUND LEVEL METER (SLM) MEASUREMENTS.

	ESSNL from Table 4 & Figure 20			SLM from PMU 6 (ref 5)		
Fire Rooms	Mean	Standard Deviation	N	Mean	Standard Deviation	N
1	90.5	3.1	17	92.0	2.8	18
2	89.8	3.9	14	88.4	3.5	17
3	87.0	0.0	1	89.5	4.2	15
4	80.0	0.0	1	91.4	2.9	14
Fwd Eng Rm	89.5	4.7	18	92.8	3.2	63
After Steering	91.1	6.5	11	94.0	6.0	37
Avg	89.9	4.5	62	91.6	7.9	165

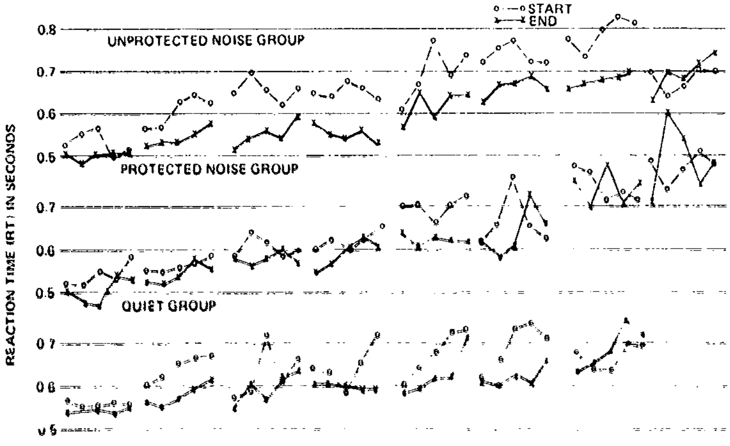

Figure 21. Subjects rank-ordered by RT on end-of-watch test.

rank-ordered from left to right on the basis of the reaction time (RT) test at the end of the watch, which in actuality is the second giving of the test or retest. Those at the extreme left have the shortest RT, and each succeeding subject to the right takes progressively longer to react. The general trend for all subjects is to deteriorate in performance (longer RTs) for each succeeding minute of the test and to increase in performance (shorter RTs) for the second test (at the end of a 4-hour watch).

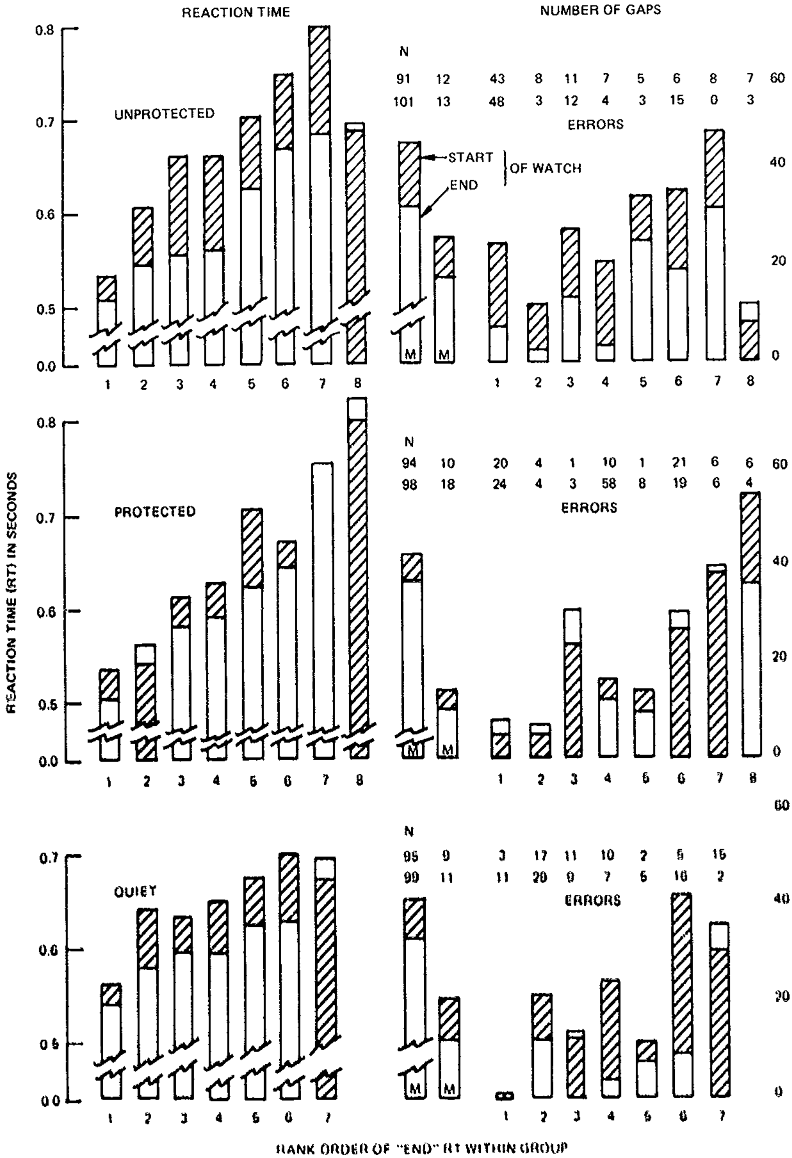

Figure 22. Average RTs, gaps, and errors (table) for each subject on start- and end-of-watch tests rank-ordered by end-of-watch RT.

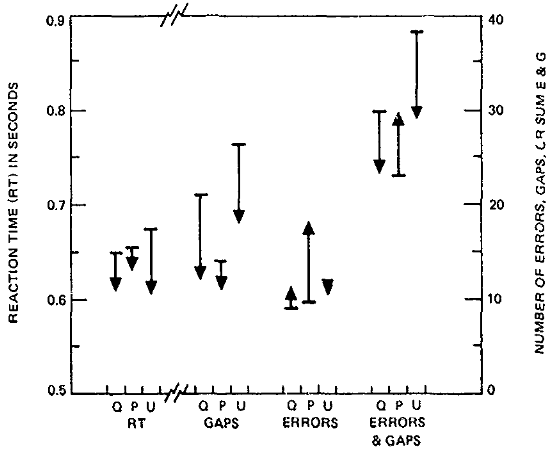

Figure 23. Average RTs, gaps, errors, and gap plus errors.

A measure of the deterioration (increase) in RT over the 5-minute testing period is a positive difference between the sum of the RTs at the end of minutes 4 and 5 and the sum of minutes 1 and 2. Table 6 shows these differences for each subject in each group and confirms that 36 of the 46 differences are positive. The interesting point is that both noise groups slowed down more after 5 minutes of testing at the end of the watch than at the beginning of the watch, while the quiet group did the reverse.

Another measure of performance decrement with prolonged testing is the increase of abnormally long response times first noted and called "blocks" by Bills (ref 8, 1931) and now often called "gaps" (see summary review by Poulton, ref 9, 1971). In this four-choice serial reaction time task, a gap is defined as any response time greater than 1 second. The difference between the number of gaps at the end of minutes 4 and 5 and of minutes 1 and 2 gives a measure of this type of deterioration. These differences are also shown in table 6. Note that the quiet subjects show a decreasing rate of deterioration on the end-of-watch test while the two noise groups show the reverse. In summary, both the RT and gap measures show greater rates of deterioration for the two noise groups on the end-of-watch test and less for the quiet group, although none of these observations reach statistical significance. It should be remembered that these subjects took the tests during their watches in their watch stations. There were noticeable job-related distractions, such as their mates' reacting in very active and meaningful ways to gauge pointers out of limits, etc. and some of these "incidents" are reflected in spurious points in figure 21. It is believed that larger samplings in slightly more remote positions would have tightened up the results. As they stand, however, they truly represent the effects of the environment on the man's performance. It is not possible to have realism and still get results with laboratory-like precision.

---

[8] Bills, AG (1931). Blocking-A new principle in mental fatigue, Am J Psych, 43, 230-245
[9] Poulton, EC (1971). Environment and Human Efficiency. Springfield, Illinois, Thomas

TABLE 6. SUMMARY OF DETERIORATION IN REACTION TIME (RT) AND INCREASE IN GAPS BETWEEN TESTS 1 AND 2 (START AND END OF WATCH). THE DATA ARE DIFFERENCES IN RT OR GAPS BETWEEN THE SUM OF THE SCORES AT THE END OF MINUTES 4 AND 5 AND THE SUM FOR MINUTES 1 AND 2.

	Reaction Time			Gaps		
	End	Start	Difference	End	Start	Difference
	3	-4	7	0	0	0
	98	116	-18	3	6	-3
	98	112	-14	4	6	-2
Quiet Group (Q)	-29	91	-120	-3	8	-11
	155	217	-62	4	2	2
	55	185	-130	2	9	-7
	191	71	120	12	7	5
Mn	81.6	112.6	-31.0	3.1	5.4	-2.3
σ	78.5	73.0	85.0	4.6	3.3	5.0
	79	69	10	4	2	2
	60	84	-24	5	4	1
	21	-50	71	3	-4	7
	121	45	76	-3	3	-6
Protected Noise Group	-3	26	-29	0	4	-4
	189	9	180	7	7	-0
(PN)	12	-127	139	7	-12	19
	-71	87	-158	7	11	-4
Mn	51.0	17.9	33.1	3.8	1.9	1.9
σ	80.2	73.9	107.1	3.7	7.0	8.1
	33	-64	97	-2	-10	8
	72	139	-67	2	4	-2
	87	-68	155	6	1	5
Unprotected Noise Group	-35	6	-41	-1	8	-9
	72	148	-76	0	-2	2
(UN)	52	-35	87	1	2	-1
	50	133	-83	5	6	-1
	142	66	76	4	-1	5
Mn	59.1	40.6	18.5	1.9	1.0	0.9
σ	50.2	92.6	94.8	2.9	5.6	5.3

In addition to ongoing changes it is also important to look at average changes for the whole 5-minute test between the start- and end-of-watch tests. These are evident in figure 21, where it can be noted that in the unprotected noise group (UN), seven of eight subjects show consistently reduced RTs and gaps at the end of the watch.

The protected group (PN) shows only 1 consistent (improved) subject, number 5, three highly variable, and four subjects that vary inconsistently and minimally between the beginning (first) and end (second) of watch tests. On the average (see fig 22), the end-of-watch RTs are slightly shorter.

The control group (Q) results in figure 21 show four subjects with consistent end-of-watch improvements and three variable subjects, but six of seven with an average improvement (fig 22). The gap data (fig 22) disagree on only one subject, number 3.

In general, the error rate is not high in this type of test, and that was true for this sample. Errors are listed in tabular form on figure 22, and there are but five subjects who make more than 20 errors (out of about 500 tries per subject, so that 20 errors represents an error rate of 4%). Three of these are in the PN group.

Figure 23 shows experimental group averages in terms of RT, number of errors, gaps, and gaps plus errors. Note that all groups show a reduction in RT and gaps, and all except the PN group show a reduction in errors plus gaps.

In summary, the four-choice serial reaction test results given within the first half-hour (start) and last half-hour (end) of a watch show that all subjects tend to deteriorate toward the last 2 minutes of the 5-minute task both at the start and end of a watch. However, the rate of deterioration for the two noise groups is greater at the end of the watch while the reverse is true for the quiet group. Averaged over all 5 minutes of the test, all groups show improved performance on the end-of-watch test, due in large measure to the fact that it was the second time they took the test. The amount of improvement was greatest for the UN group, who also did the worst on the first test. The improvement in RT (shorter) and tries (more) was accompanied by a reduction in gaps.

The results could be explained by assuming the noise aroused the UN group to greater effort (more tries and shorter RTs), but that within the 5-minute sustained test the rate of deterioration was increased by the noise exposure.

MODIFIED STROOP COLOR WORD TESTS

Since the subject populations in these tests kept changing, it was not feasible to run this test on subjects prior to the time-in-noise at sea. Therefore, since there was no baseline test, there was a learning factor involved and the second test (given at least 3 hours after the first) was always performed in less time, as can be seen in table 7 and figures 24 and 25. Table 7 and figures 24 and 25 show the results averaged over the three groups (Q, PN, and UN) for the three types of Stroop materials (WORD, HUE, and MIXED) for both the start-of-watch (1) time reduction ratios (fig 24) and (2) differences between all combinations of WORD, HUE, and MIXED TIMES (fig 25). Arrow lengths represent changes in these quantities with the second giving of the test. Note that: (1) all groups performed all tasks in less time on the second presentation of the test, (2) all groups took less time to sort the HUE cards and more time to sort the MIXED cards, and (3) the Q group took appreciably more time on the first MIXED sort and showed the greatest improvement on the second MIXED sort. The amount of improvement (time reduction) on the end-of-watch test is plotted at the bottom of figure 24. Over all groups and materials the time reduction ratio (T_2/T_1) averaged 0.81 (left ordinate) so the average improvement was 1.23 (right ordinate).

The significance of the Stroop test lies in the amount of perceptual interference induced by printing the color names in incompatible hues here expressed as the difference between the MIXED score and the HUE score and/or the difference between the MIXED and WORD scores. The HUE minus WORD differential is a neutral measure that can be used as a baseline or comparison measure for the other two. The three differences for both

TABLE 7. SUMMARY OF MODIFIED STROOP COLOR TEST RESULTS IN SECONDS (WITH CORRECTION FOR ERRORS; IE, 3 SECONDS ADDED PER ERROR).

		W	H	M	H-W	M-W	M-H	
Quiet	1	101	92	136	-9.0	35.7	44.7	1
Group	2	82	73	92	-8.6	10.1	18.7	2
n = 9	2/1	0.83	0.80	0.68	-0.4	25.6	26.0	1-2
Protected	1	97	85	125	-12.1	29.5	38.9	1
Noise	2	82	76	96	-6.2	13.5	19.3	2
Group	2/1	0.85	0.85	0.77	-5.9	16.0	19.6	1-2
n = 13								
Unprotected	1	105	93	125	-12.0	19.9	31.9	1
Noise	2	83	77	101	-5.5	17.6	23.2	2
Group	2/1	0.79	0.83	0.81	-6.5	2.3	8.7	1-2
n = 13								

W = Pairs of color name WORDS in black ink
H = Pairs of color (HUE) patches
M = Pairs of color name words in incompatible ink hues (MIXED)
1 = Start-of-Watch test (within first half-hour of 4-hour watch)
2 = End-of-Watch test (within last half-hour of 4-hour watch)
2/1 Ratio of End-of-Watch to Start-of-Watch (time reduction ratio)
H-W, M-W, M-H equal differences in sorting time scores

the first and second test on each of the three groups are shown in figure 25. Both measures of perceptual interference show an appreciable reduction for the two control (quiet and protected noise) groups between the first and second test. Analyses of variance tasks on these differences fail to reach significance, which is not surprising considering that the tasks were given on watch with no prewatch or baseline practice test and the ongoing environmental distractions were great. It is believed a larger sample given in situ but 20 feet away (with back turned) from their usual watch station would increase the precision of measurement.

The reasons for the trends shown in the Stroop test results are far from clear. For one thing, the unprotected group showed far less perceptual interference on the first attempt at the test (after being in 94(±7) dBA for 15 to 45 minutes). This could be an artifact from the way the groups were (inadvertently) chosen; to wit: it became apparent after the initial assignment of noise-exposed men to the protected and the unprotected groups that many of the unprotected actually were sound-powered phone (SPP) talkers on at least some watches and therefore wore an earphone-in-muff -a type of hearing protection. These men were considered to be in the protected group if they were SPP talkers when given the Stroop and reaction time tests. However, they may have been selected as SPP talkers because of their linguistic abilities and the perceptual interference of the Stroop test is a linguistic interference. At any rate, it should be remembered that at most these men represented fewer than half the people in the protected noise group.

Another source of confusion in interpreting Stroop color word interference in noise is the presence or absence of an auditory/verbal factor. Do subjects sub-vocalize the name of the color in the process of making their decisions? There was no reason to expect that

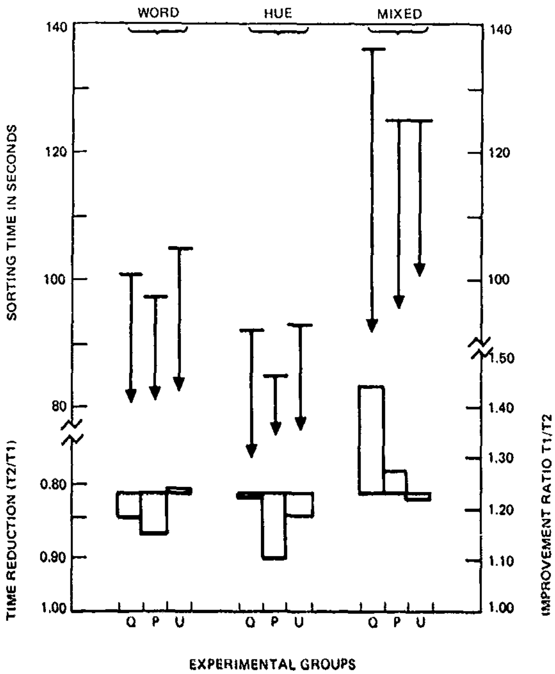

Figure 24. Sorting times and improvement ratios for the three types of modified Stroop color word test (SCWT) materials (WORD, HUE, MIXED) at start and end of watches.

they did, since the sorting bins were labeled by HUE (patches of color), not by WORD. However, there is no way of knowing what subconscious cues the subjects did in fact use.

In any case, the use of WORD, HUE, or MIXED pairs as opposed to a single WORD, HUE, or MIXED word makes this modification of the Stroop test different from its usual application. Therefore, comparing these results to any other Stroop/noise results is of dubious value. The results do, however, agree in kind with the results of the second experiment described by Hartley and Adams (ref 2, 1974). Their "watch" period was only 30 minutes as compared to our 240 minutes and therefore their start-of-watch test was completed within the first 10 minutes, whereas their end-of-watch test was the last 10 minutes of a 30-minute period. Their end-of-watch test was probably more similar to our start-of-watch test, since it was often 20 to 40 minutes into a watch before we completed our first test. Nevertheless, our major findings are similar in that the "noise" group did far worse on the MIXED (or their experimental) test at the start of the watch than did the quiet group. And, similarly, both their and our noise groups showed little or no change in

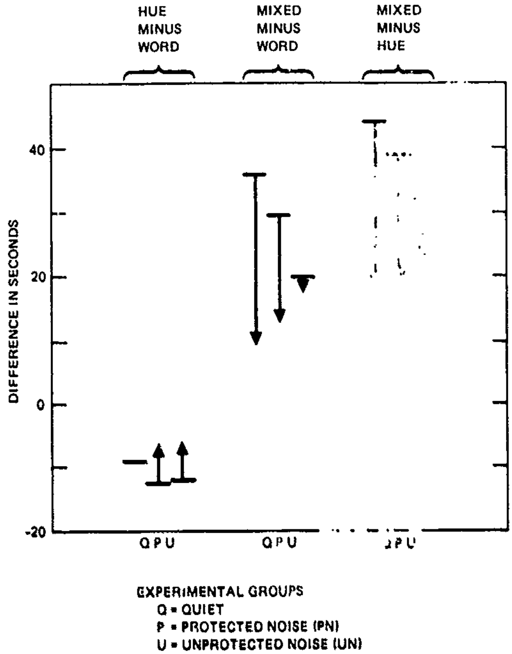

Figure 25. Differences in sorting times for each pair of SCWT materials (HUE-WORD, MIXED-WORD, MIXED-HUE).

interference scores at the end of the watch as compared to the quiet groups who, in both cases, improved (showed less perceptual interference) at the end of the watch.

CONCLUSIONS

The major observation of this shipboard (field) study is the almost insurmountable difficulty of running a controlled experiment on a not-to-interfere basis on subjects who are not highly motivated. Although all subjects did indeed sign "volunteer" statements, they were not true volunteers. That is, it could not be said that there were available more than enough subjects of the type needed and only the eager, conscientious ones stepped forth. Their watch schedule was altered to include a "dog watch" every day vice once every fifth day or week, and this seriously affected their sleeping routine—at least, such complaints were stated. The experimenters had been led to believe that a daily dog watch was typical in the 1-in-3 routine and were not prepared for the antagonism this created.

A second related observation was that the watch cannot be staggered by a factor of ±½–¾ hour, even when preplanned and agreed to (in principle) with both the officers in charge and the three-man teams assigned to expedite it. It appears that if extra duty for which the subject sees no gain to himself is added, it must be exacted in a hard-nosed, chain-of-command, continually supervised effort. This reaction in itself may say something about the fatiguing, irritating aspects of standing 4-hour watches in hot, humid, noisy engine rooms.

Since it turned out to be impossible to get adequate immediate postwatch (noise) audiograms, the compromise of finding 10-day (vice 4-hour) semipermanent threshold shifts was used to evaluate the effects of continual 1-in-3 watches (essentially 4 hours in noise and 8 hours out) over a 10-day period. In this manner, eight (of 39) men were found with observable threshold shifts from hearing performance levels shown by their baseline audiograms (taken before the ship left port and after being in port for a minimum of 2 weeks; ie, the subjects not standing watches in the noisy fire rooms). Of the eight, one was from the control, or work-in-quiet (the Engineering Log Room), group of 12 subjects. This subject did, however, listen to very loud "music" whenever he had any leisure time. Of the remaining seven, all stood watches in noisy areas (94 ± 7 dBA); four on no occasion wore any hearing protection; three wore no protection but on some occasions wore SPP earphones in cushions. These seven were from a total sample of 23 noise-exposed unprotected subjects. None of the 11 subjects who always wore hearing protection were among the threshold shift group.

Dosimeter results showed that the typical noise exposure in terms of equivalent steady-state noise level in the space where most measurements were made averaged 89.9 dBA with a standard deviation of 4.5, whereas physical measures averaged 91.6 dBA with a standard deviation of 7.9.

On a psychomotor four-choice serial reaction time test, all subjects (in noise and quiet) deteriorated on the average when performance at minutes 4 and 5 (of a 5-minute test) was compared to performance the first 2 minutes. Although not statistically significant, the quiet group deteriorated less than the protected or unprotected noise group. Similarly, all three groups showed improvement in performance on the reaction time test near the end of the watch as compared to performance near the start of the watch. There was apparently a learning factor. The unprotected noise group showed the greatest end-of-watch improvement, but this could have been because they performed worse than the others at the start of the watch (after from 10 to 30 minutes in noise).

Perceptual interference was measured by a modification of the Stroop color word test. Again, all groups improved their overall performance on the retest toward the end of a watch. The unprotected noise group had the least perceptual interference on the start-of-watch test and the quiet group the most interference. Both control groups (quiet and protected noise) showed reductions in perceptual interference at the end of a watch; the unprotected noise group did not.

RECOMMENDATIONS

1. Verify the results of this preliminary study by repeating the most significant parts with an expanded shipboard team including a USN Chief Petty Officer whose responsibility would be to expedite and achieve compliance with the controlled testing schedules.

2. Develop a wearable noise level accumulator to supplement existing noise dosimeters (too simple) and to better correlate performance and hearing tests to actual noise exposure.

3. Implement educational and compliance techniques to get more people wearing hearing protection to help save their hearing, improve their psychomotor performance, and reduce their perceptual interference in noise.

REFERENCES

1. Cantrell, RW (1974), Prolonged exposure to intermittent noise: Audiometric, Biochemical, Motor, Psychological and Sleep Effects, Laryngoscope, Supplement 1, 84 (10), Part 2, October

2. Hartley, LR, and Adams, RG (1974), Effect of noise on the Stroop test, J Exper Psych. 102 (1), 62-66

3. US Department of Labor, Occupational Safety and Health Standards, Federal Register, XXXVII, Part II, Washington, DC, 1972

4. US Navy, Bureau of Medicine Instruction 6260.6B, 5 March 1970

5. US Navy, Environmental and Preventive Medicine Unit 6 Encl (2) to ltr Ser EP MU6/6: vn, 6265 of 8 August 1975 to USS ORISKANY

6. Wilkinson, RT (1969), Some factors influencing the effect of environmental stresses on performance, Psych Bull 72, 260-272

7. Wilkinson, RT, and Houghton, D (1975), Portable Four-Choice Reaction Time Test with Magnetic Tape Memory, Behavior Research Methods and Instrumentation, 7 (5), 441-446, September

8. Bills, AG (1931), Blocking--A new principle in mental fatigue, Am J Psych, 43, 230-245

9. Poulton, EC (1971), Environment and Human Efficiency, Springfield, Illinois, Thomas